Consumer Complaints

**ExcelR**

**US Data**

<https://catalog.data.gov/dataset/consumer-complaint-database>

# Task

Given a consumer complaint, predict which product is complaint about. This is a Multi-Class text classification.

# Outcome

Product – This column in the dataset needs to be considered as Outcome/Label.

# Description

This data set is from US Government and our goal is to use sequence learners [RNN/LSTM] and try classifying it to one of the classes from column “Product”. Given that this is a text classification problem and based on our internal analysis, we have realized that this task can be well achieved using column “Consumer Complaint Narrative”, but if required participants are free to explore other columns.

# Column Details

<https://catalog.data.gov/dataset/consumer-complaint-database>

# Hint

Classes are imbalanced, so treating that is going to be important. As we are dealing with text data which is sequential in nature it’s advisable to use sequential models like RNN/LSTM, but participants are free to explore.